

AMENDMENTS TO THE CLAIMS

1-2. (CANCELLED)

3. (CURRENTLY AMENDED) A computer-implemented method of constructing a model for predicting molecular behavior using marker molecules, said method comprising:

classifying respective molecules in an training set of reference molecules as either possessing or not possessing at least one chemical or biological property;

selecting a first subset of said training set of reference molecules, wherein all of the molecules in said subset possess the at least one property;

comparing all molecules in said training set with all other molecules in said training set in accordance with a pre-defined ~~numerical-structural~~ similarity metric;

selecting a target molecule from said first subset;

defining, for a first each other molecule in said training set other than said target molecule, a fractions-correctly-predicted metric as the number of molecules in said training set that are members of said first subset and that have a ~~numerical-structural~~ similarity to said target molecule at least as great as said ~~each other~~first molecule's ~~numerical-structural~~ similarity to said a-target molecule divided by the total number of molecules in said training set having a ~~numerical-structural~~ similarity to said target molecule at least as great as said each other molecule's ~~numerical-structural~~ similarity to said target molecule;

repeating the defining step for each molecule in said training set other than said target molecule;

determining, from molecules in said training set having a fractions-correctly-predicted metric below a threshold value, which molecule has the highest ~~numerical structural~~ similarity to said target molecule;

counting the number of molecules in said training set having a higher ~~numerical structural~~ similarity to said target molecule than said molecule determined in said determining step;

choosing said target molecule as a marker molecule if said number of molecules determined in the counting step is equal to or greater than a pre-selected value-; and

outputting data indicating that said target molecule has been chosen as a marker molecule.

4. (CURRENTLY AMENDED) The method of Claim 3, wherein said first subset comprises all of the molecules in said training set that possess said at least one property.

5. (CANCELED)

6. (PREVIOUSLY PRESENTED) The method of Claim 3, additionally comprising repeating said determining and counting steps for a plurality of different threshold values.

7. (CURRENTLY AMENDED) The method of Claim 3, comprising repeating said selecting a target molecule, defining, repeating, determining, counting, and choosing steps for other molecules of said first subset at a plurality of different threshold values and pre-selected number of molecules value so as to select a plurality of preliminary sets of marker molecules.

8. (CURRENTLY AMENDED) The method of Claim 7, comprising choosing a final set of marker molecules by making molecular behavior predictions for all molecules in said training set using each one of said preliminary sets of marker molecules, and choosing as said final set of marker molecules the preliminary set that most accurately predicts molecular behavior of molecules of said training set.

9-18. (CANCELED)